STATE INNOVATION MODEL INITIATIVE HEALTH INFORMATION TECHNOLOGY COUNCIL

HEALTH INFORMATION TECHNOLOGY (HIT) BRIEFING

DECEMBER 18, 2014

Roderick L. Bremby

Commissioner, Department of Social Services

HIT Council, Co-Chair

Briefing Outline

- SIM HIT Plan Overview – 20 Minutes
  - Review HIT Plan components by function
- Level Set Using Analogy – 5 Minutes
- Discuss Conceptual Readiness – 10 Minutes
  - Logic Model
- Next Steps – HIT Council Function – 10 Minutes
- Q & A

Objectives

- To review SIM HIT plan components
- To demystify plan technology – Analogize A Journey to Well-being
- To discuss next steps in the evolution of the plan.

Briefing Outline

- SIM HIT Plan Overview – 20 Minutes
  - Review HIT Plan components by function
- Level Set Using Analogy – 5 Minutes
- Discuss Conceptual Readiness – 10 Minutes
  - Logic Model
- Next Steps – HIT Council Function – 10 Minutes
- Q & A

HIT Plan Technical Components

- # *CareAnalyzer*® (DST Healthcare Solutions)
- # HIT Strategic Plan
- # APCD – All Payers Claims Database
- **Consent Registry -**
- **Disease Registries -**
- Crowd Sourcing- method of engaging large groups to contribute to an outcome
- # EMPI - enterprise Master Patient Index (Nextgate)
- **Personal Health Record – MyChart (Epic)**
- # Provider Directory- (NextGate)
- # Direct Messaging/Admission, Discharge, Transfer - Secure Exchange
- **Edge Servers/Indexing/eCQM (e Clinical Quality Measures)**
- Electronic Health Records Service as Software Model

 # Current assets in process or production

SIM HIT Plan - Proposed initiatives

- HIT Governance Structure
- Consent Registry
- Disease Registries
- eCQMs reporting engine using edge servers
- Alert/Notification Engine
- Personal Health Record

HIT Governance

- Starting on October 15, 2014 a set of six-planning meetings with a focus on the following:
  - Create a HIT vision statement for our state
  - Identify common HIT goals
  - Identify and support an enterprise built on an interoperability framework
  - Operationalize across-agency governance structure that builds upon and ties the various initiatives that have been undertaken in the last 4-years with respect to health and human services.

SIM HIT Plan Budget

SIM HIT Budget - Request Grant & Bond Funding		State	Federal
Personnel (5.5 FTE)	Research Associate 2		**2,245,269**
Travel/Supplies	Travel (Instate & NE Region/Mileage)		80,000
	Printing/Publishing educational materials		12,000
	Computers and software		10,000
	Office Supplies		6,000
	Subtotal		**108,000**
Procurement/Contractor Cost	BEST Hosting EMPI and PD		480,000
	Care Analyzer		700,000
	HIT Strategic Plan		200,000
	APCD		540,000
	Consent Registry	900,000	1,100,000
	Disease Registries		2,200,000
	Crowd Sourcing		360,000
	EMPI-Nextgate		208,600
	Provider Directory- NextGate		225,000
	Direct Messaging/ADT - Secure Exchange		450,000
	Edge Servers/Indexing/eCQM	900,000	1,000,000
	EHRs SAAS Model		735,000
	Subtotal		**8,198,600**
Indirect			**233,727**
Grand Total		**1,800,000**	**10,769,596**

Briefing Outline

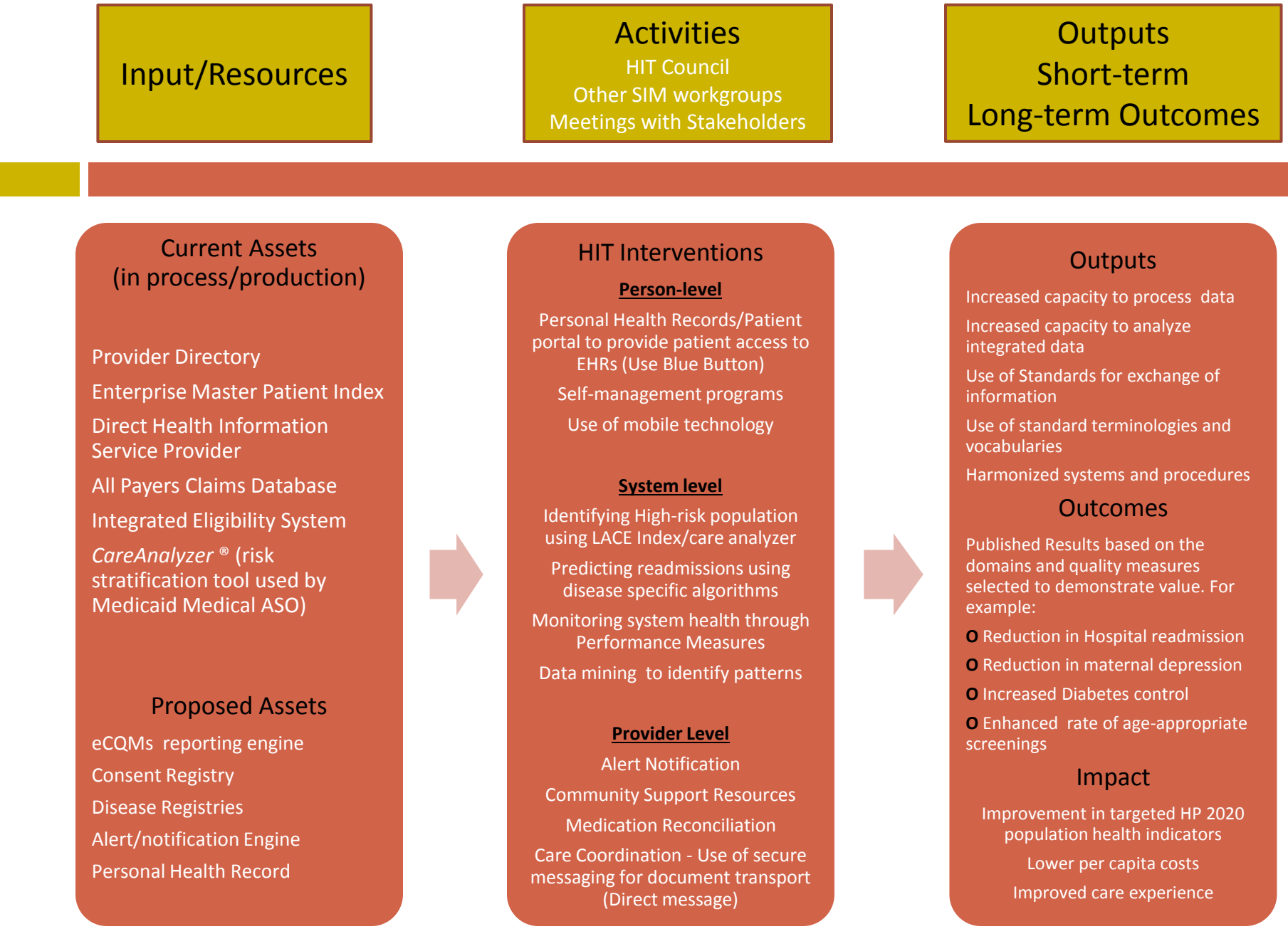
- SIM HIT Plan Overview – 20 Minutes
  - Review HIT Plan components by function
- Level Set Using Analogy – 5 Minutes
- Discuss Conceptual Readiness – 10 Minutes
  - Logic Model
- Next Steps – HIT Council Function – 5 Minutes
- Q & A

A Journey to Well-being

When you get the what, you know the how. – Thomas Horst

Briefing Outline

- SIM HIT Plan Overview – 20 Minutes
  - Review HIT Plan components by function
- Level Set Using Analogy – 5 Minutes
- Discuss Conceptual Readiness – 10 Minutes
  - Logic Model
- Next Steps – HIT Council Function – 10 Minutes
- Q & A


Input/Resources
Activities
HIT Council
Other SIM workgroups
Meetings with Stakeholders
Outputs
Short-term
Long-term Outcomes
Current Assets
(in process/production)
Provider Directory
Enterprise Master Patient Index
Direct Health Information Service Provider
All Payers Claims Database
Integrated Eligibility System
CareAnalyzer ® (risk stratification tool used by Medicaid Medical ASO)
Proposed Assets
eCQMs reporting engine
Consent Registry
Disease Registries
Alert/notification Engine
Personal Health Record
HIT Interventions
Person-level
Personal Health Records/Patient portal to provide patient access to EHRs (Use Blue Button)
Self-management programs
Use of mobile technology
System level
Identifying High-risk population using LACE Index/care analyzer
Predicting readmissions using disease specific algorithms
Monitoring system health through Performance Measures
Data mining to identify patterns
Provider Level
Alert Notification
Community Support Resources
Medication Reconciliation
Care Coordination - Use of secure messaging for document transport (Direct message)
Outputs
Increased capacity to process data
Increased capacity to analyze integrated data
Use of Standards for exchange of information
Use of standard terminologies and vocabularies
Harmonized systems and procedures
Outcomes
Published Results based on the domains and quality measures selected to demonstrate value. For example:
O Reduction in Hospital readmission
O Reduction in maternal depression
O Increased Diabetes control
O Enhanced rate of age-appropriate screenings
Impact
Improvement in targeted HP 2020 population health indicators
Lower per capita costs
Improved care experience

Briefing Outline

- SIM HIT Plan Overview – 20 Minutes
  - Review HIT Plan components by function
- Level Set Using Analogy – 5 Minutes
- Discuss Conceptual Readiness – 10 Minutes
  - Logic Model
- Next Steps – HIT Council Function – 10 Minutes
- Q & A

HIT Council Function

- Representatives
- Meeting frequency
- Charter
- How the work will be performed
- Recommendations to the SIM Steering Committee

Briefing Outline

- SIM HIT Plan Overview – 20 Minutes
  - Review HIT Plan components by function
- Level Set Using Analogy – 5 Minutes
- Discuss Conceptual Readiness – 10 Minutes
  - Logic Model
- Next Steps – HIT Council Function – 10 Minutes
- Q & A

Zato Health Interoperability Platform

Secure Federated Analysis Across Data Silos

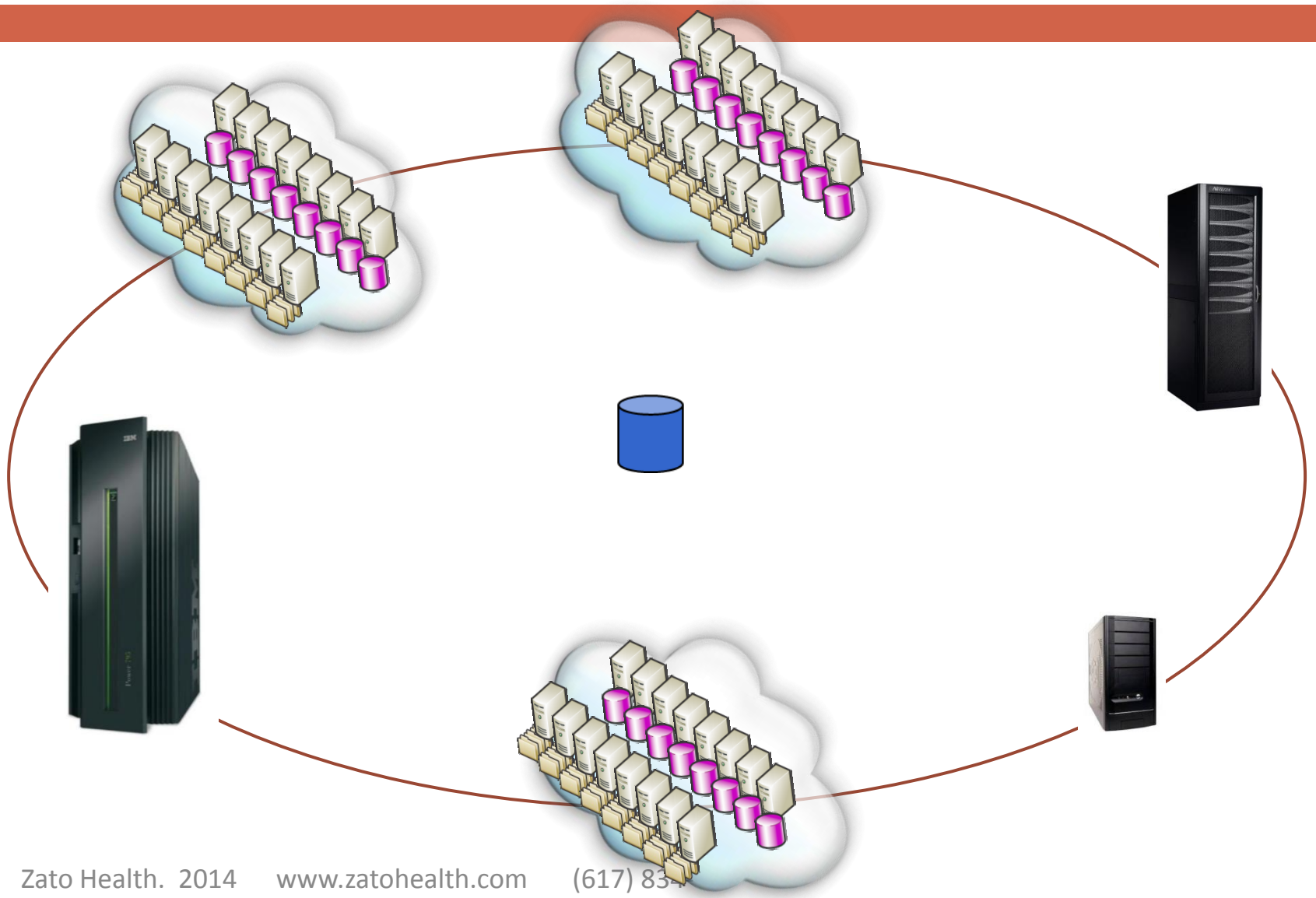
Cooperative Computing at the Edge with Cross-Network Information Fusion

December, 2014

Stove Piped Data Silos
Barriers to Improved Productivity and Cost Effectiveness

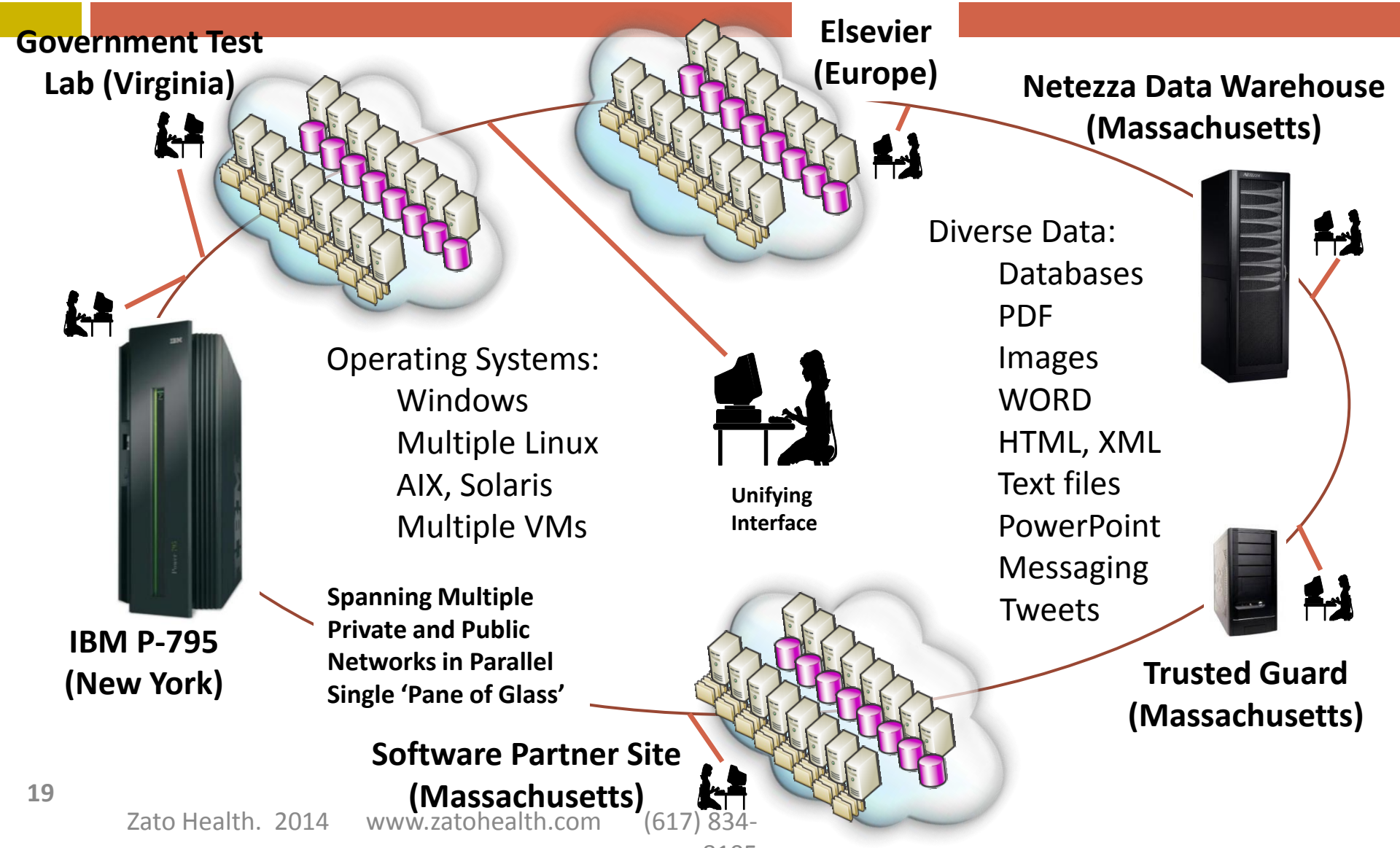
Today's Primary Data Analysis Approach

Data Aggregation and Centralization: Larger Stove Piped Data Silos



Secure Cross-Network Analysis and Discovery in Parallel in Seconds: 100 Data Sources, 500 Terabytes, 32 Billion Records & Documents

IC Sponsored World-Wide Performance Demonstration of Zato Component Technology



Proven Capabilities to Span the Following Stovepipe Barriers in Parallel at Massive Scale with a Single User Action

1. 'Structured' tabular data and 'unstructured' text-rich medical documents in EHRs and other clinical and genomic databases and applications (data silos)
2. Multiple organizations with multiple decentralized data centers
3. Multiple private and public networks and organizational security boundaries
4. Multiple file 'name spaces,' table spaces, and distributed file systems
5. Data at rest and newly arriving data in streaming messages, updates, feeds
6. Multiple hardware platforms (servers, clusters, appliances, mainframes, clouds)
7. Multiple operating systems with and without different virtual machines (VMs)

Differentiated Performance Characteristics for Cross-Organization Information Sharing

- **Massive Scalability** across decentralized application repositories, with capacity for scores of petabytes, trillions of records and documents
- **Consistent global relevance ranking accuracy as** data volumes increase and number of nodes increases across data centers
- **Consistent response time performance** as data volumes increase and number of nodes increases across data centers
- **Efficiency over existing networks** across multiple data centers
- **Cross-Domain access control security** executed at the edge
- **Unobtrusive** to local production information systems during query and analysis across production data repositories

Zato Principals
(www.zatohealth.com)

- Paul McOwen (CEO) – Administrative Director of a National Science Foundation research center; Deputy Chairman of Computer Science Department at University of Massachusetts, Amherst; Founder and CEO of high growth software companies with disruptive technology

- Dr. Daniel Heinze, PhD (Founder and Chief Scientist) – Lead developer for DARPA programs; inventor/architect for Computer Assisted Coding systems, automated text processing; Founder, CTO of A-Life Medical (acquired by UnitedHealth in 2010, now OptumInsight™)

- Dr. John Holbrook, MD (Founder) – Created first 24/7 Hospitalist program in the U.S., Director, Chief Medical Officer, VP, and consultant for hospitals, hospital associations, state programs, disruptive technology companies and leading insurance companies

- Dr. Winthrop F. Whitcomb, MD (Founder) – Co-Founder of the Society of Hospital Medicine (> 10,000 healthcare professionals), Medical Director and Chief Medical Officer for hospitals and for innovative healthcare software companies

Accomplishments of the Zato Team

- **Created the most widely used Computer Assisted Coding (CAC) system for automated medical text processing and coding in the healthcare industry**
- Created comprehensive clinical ontology and advanced techniques for medical natural language processing and seamlessly integrated both
- **Created information fusion for U.S. intelligence and defense agencies with extraordinary customer testimonials for productivity gains and outcomes**
- Field proven capability to securely span stovepipe boundaries with one user action from a 'single pane of glass' across massive diverse data silos

Barriers to Healthcare Interoperability and Data Liquidity

- Costs and risks of cross-organization data export (loss of control, accountability)
- A ‘Tower of Babel’ of diverse and complex EHRs, other database applications, full-text in medical records, internal/external centralized and de-centralized data
- High expense and limited flexibility to produce new or modified quality of care and ‘meaningful use’ reports from complex data structures in existing systems
- Slow pace of integration of advanced technologies to solve unmet need to query and analyze data across data multiple data systems (silos)

Zato Health Selected by Connecticut for State-Wide Initiatives

- 8 initial healthcare provider sites, expanding state-wide
- Access standardized provider data for Electronic Clinical Quality Measures
- Collect Meaningful Use measures related to Medicaid EHR incentive programs
- Minimize moving data
- Ensure timely access to data for reporting and audits
- Use of edge servers to eliminate the need to physically send data to the state
- Analysis spanning multiple state agency databases in parallel

2013 Webcast Panel with the Chief Medical Officer of Centers for Medicare and Medicaid and Zato Principals

The panel for this international Webcast consisted of 3 physicians:

- Dr. Patrick Conway, MD, MSc, SFHM, Chief Medical Officer (CMO) and Deputy Administrator for Innovation and Quality for the Centers for Medicare and Medicaid Services (CMS) and Director, Center for Clinical Standards and Quality (Zato is viewed as a 'Data Intermediary')

- Dr. John Holbrook, MD, MA, FACEP, CEO of Zato

- Dr. Winthrop F. Whitcomb, MD, MHM, CMO of Zato

Panel Title:

" Healthcare Reform, Quality Reporting, and Technology: How Providers and Payers Will Demonstrate Value for Patients and Populations"

Zato Application Selected by IBM in October 2014

Zato's 'DRG Dashboard' application was selected by IBM and announced at the IBM Enterprise 2014 conference and on the IBM Web site as one of the top 5 applications evaluated in an international 'application throw-down,' based on:

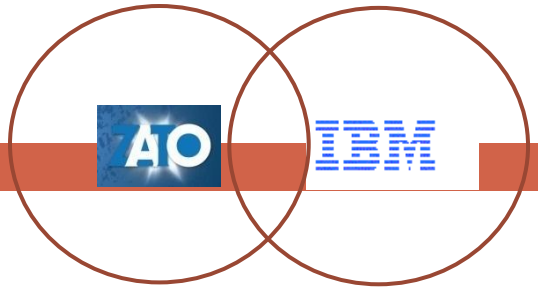
- Depth of integration
- Use of Business Analytics
- Use of Information Management
- Use if Enterprise Content Management Technology
- Level of innovation
- Demonstrable business value.

See a video of a live demonstration at http://ibmappthrowdown.tumblr.com

Benefits from Edge Analysis and Cross-Network Information Fusion

- Data interoperability across applications on servers, clusters, clouds, mainframes
- Improved discoverability from extended data reach for clinical or research uses
- Measurable productivity gains: unified views from single user action
- More affordable and flexible delivery of clinical performance reports
- Measurable improvements in quality of care, outcomes, cost effectiveness
- Enabling new clinical applications spanning clinical and genomic data
- Cost effective accountability from automated remote audits
- Greater leveraging of existing data and information system investments
- Creation of secure Healthcare Information Sharing Environments (HISE)

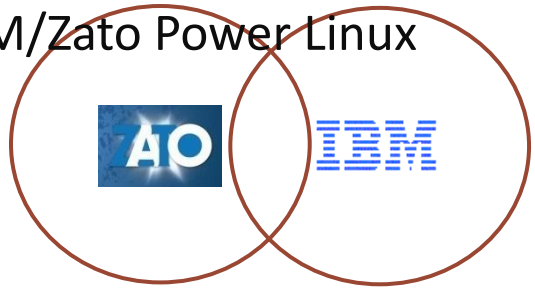
Zato Principals and IBM: A History of Cooperative Integration



- Early integrations with companies later acquired by IBM: Lotus, Language Analysis Systems (LAS), Initiate Systems, I2 Analyst Notebook, Netezza
- Performance testing of IBM/Zato P795 appliance with 500 TB (> 30 billion recs)
- Recent Solution Provider Agreement with Mainline Information Systems
- New, multi-year IBM/Zato co-marketing programs

Greater Scalability Across De-Centralized Data Will Require 'Bringing the Computation to the Data' and 'Edge Processing'

- Joint initiative with IBM Life Sciences across clinical and genomic data
- Zato Interoperability Platform testing in IBM Innovation Center
- Zato Interoperability Platform testing in IBM Poughkeepsie Research Labs
- Zato Interoperability Platform testing in the IBM Development Cloud
- 75% Zato software discounts exclusively for IBM/Zato Power Linux appliances ('IZPower'), including P8



Real-Time Functionality Across Data Silos

- Data transparency - normalizes and indexes data sources ‘at the edge’
- Interactive query and analysis - unified interface (‘single pane of glass’)
- Batch analysis and reporting with flexible output to reports or other applications
- Automatic concept recognition and entity extraction ‘on the fly’ in context
- Global navigation in parallel across data sources by context, concepts, facets
- Current awareness from continuous monitoring, real-time filtering, alerting
- Dynamic iterative knowledge discovery and visualization across data sources
- Seamless system integration and interoperability with APIs and Web interfaces
- Integrated medical ontology, automated medical coding, and medical NLP

Zato Panel at American Medical Informatics Association (AMIA) Joint Summit in San Francisco April 2014

- **"Building a Richly Connected and Highly Analyzed Genotype/Phenotype Ecosystem in a World of Data Silos"**
- Panelists:
  - CIO of Illumina "#1 Smartest Company" – MIT Technology Review
  - Director of Genomic Informatics at the Medical College of Wisconsin
  - Chief Scientist of Zato Health
  - CEO of Zato Health

Zato: Unique Integration of Two Proven Technology Innovations

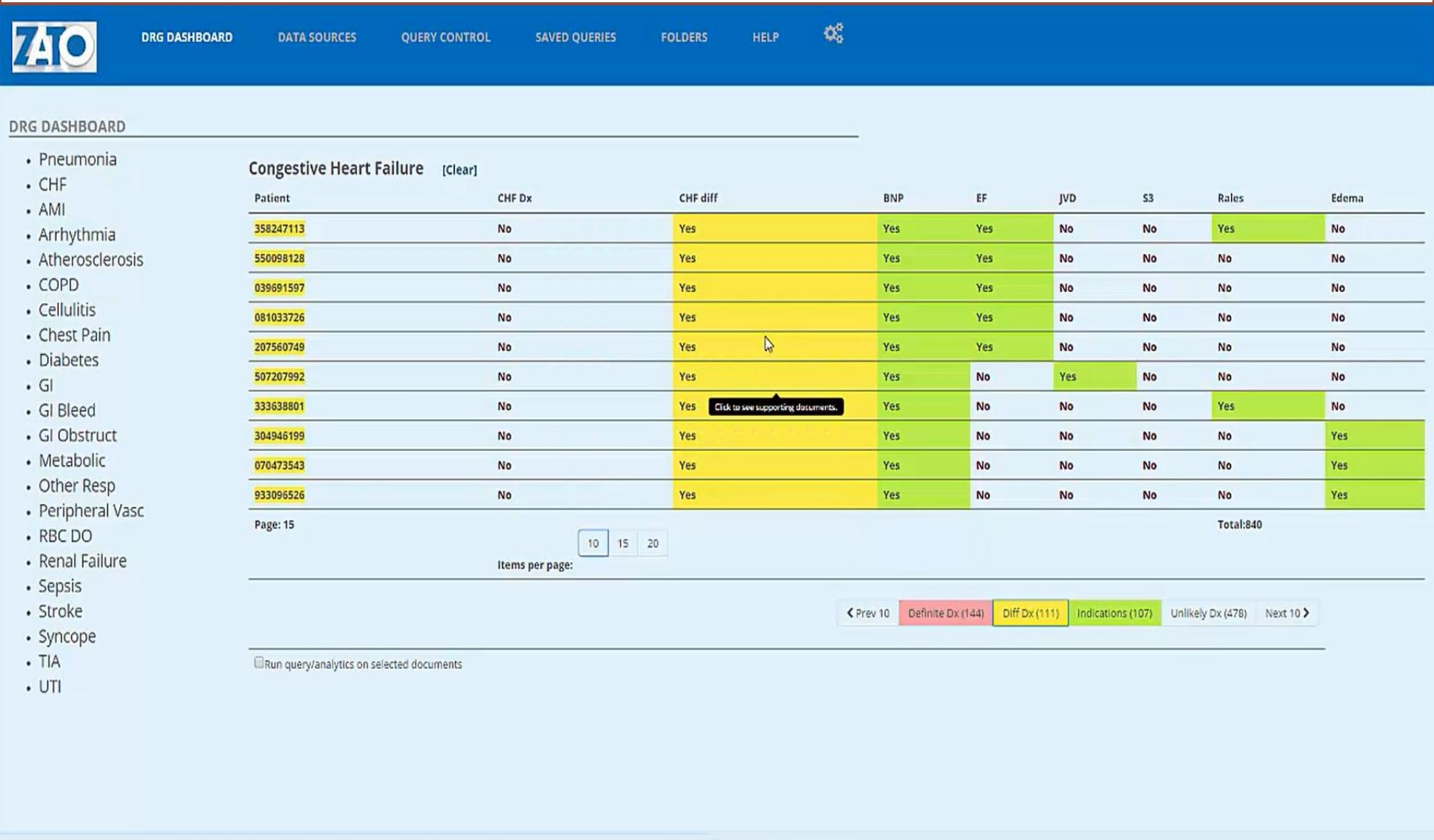
1. **A fully distributed software platform for secure information sharing:** cooperative parallel processing across diverse application data repositories in multiple data centers and organizations with real-time information fusion over multiple networks simultaneously

2. **Industry leading automated coding and advanced natural language processing for medical text** leveraging a seamlessly integrated comprehensive medical ontology

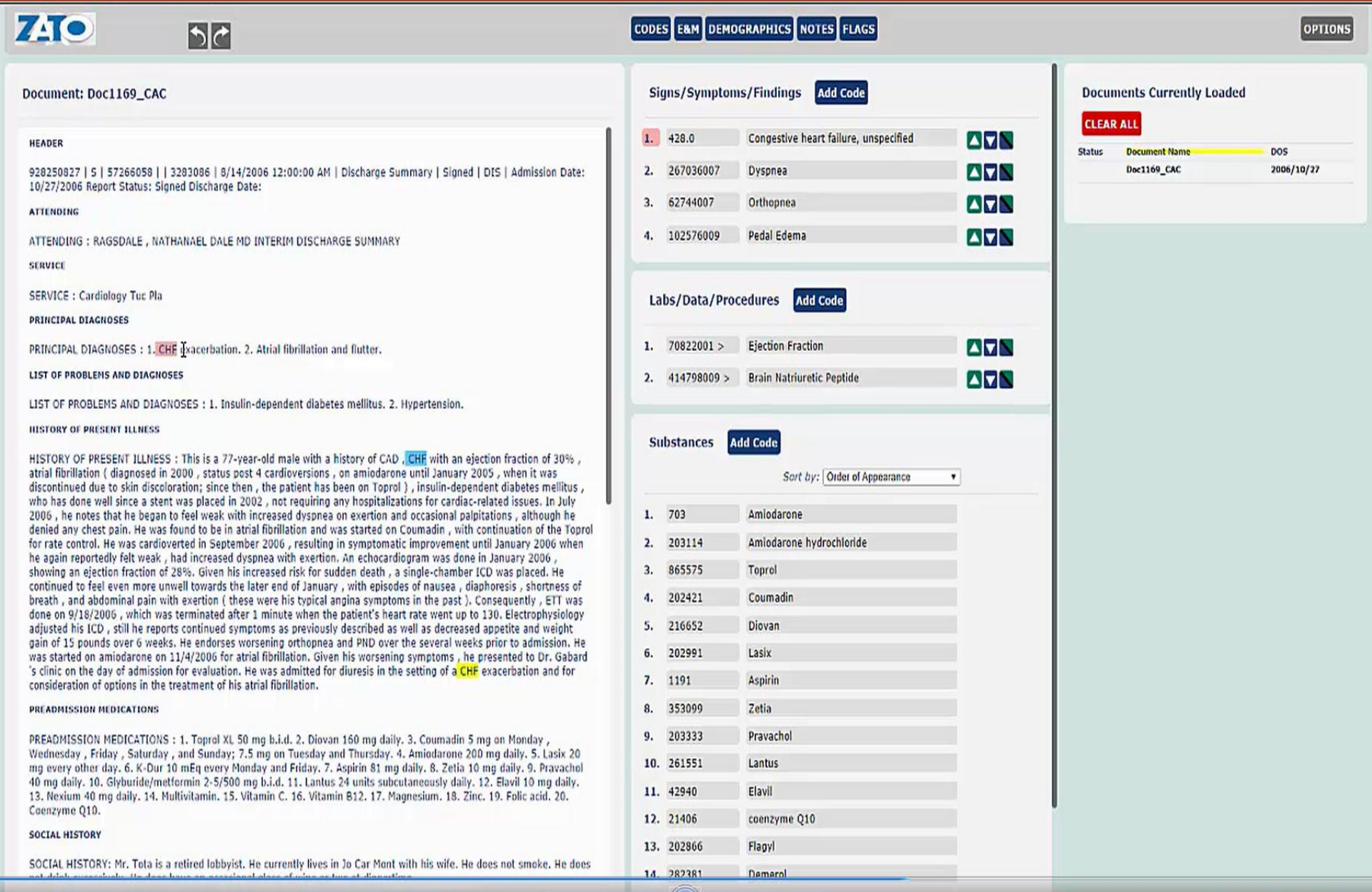
Reporting and Auditing of Continuous Improvement Measures for Intervention and Cost Effectiveness

- Automated analysis of risk across patient records for timely intervention
- Increased flexibility and reduced costs to create/modify standardized performance reports required by healthcare service payers:
  - Diagnostic Related Groups (DRGs) for Bundled Payments and intervention
  - Meaningful Use (MU) for required reporting to payers
  - Quality of care performance measures to improve patient outcomes
- Fully automated for higher productivity with verification and audit support

DRG Dashboard: Productivity for Bundled Payments and Intervention



DRG Dashboard: Example of Instant Detail from Linked Analysis

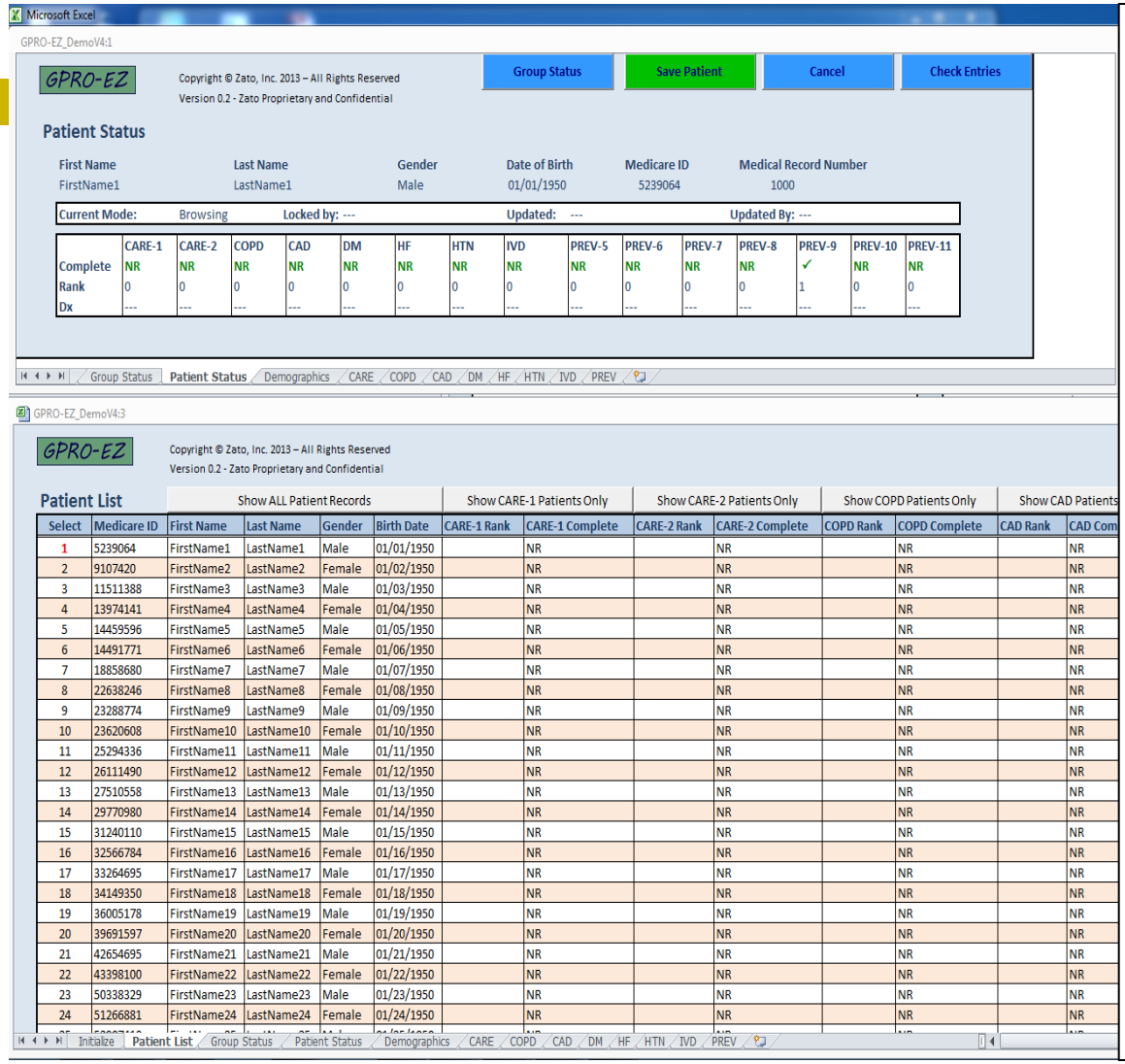


GPRO-EZ : Automated Quality Reporting for the ACO

Automated Extraction, Reporting, & Auditing of Quality Measures for Standardized Comparisons by CMS

- Bundled with the platform as an enhancement – no additional licensing fee
- Normalizes data across multiple operating systems, databases, and EHRs
- Virtualizes billions of records
- Enables semantic interoperability
- Specialized medical ontology extractors for quality of care
- Open framework for customizing & adding medical information extractions & information retrieval definitions
- Automates the quality of care information extraction and reporting from structured and unstructured data
  - Across providers, clinics and partner organizations
  - Any volume, any format, any location,
  - Without the need to move data
  - Without placing a load on existing production database systems
- Uses Zato's semantics enhanced ontology
- Populates GPRO compliant report and transmits the data to CMS

GPRO-EZ™ Electronic Reporting to CMS



The software automates the process of reporting of patient quality of care measures by ACOs to the CMS.

- **GPRO-EZ is like an automated version of an EZ form for tax preparation**
- **GPRO-EZ uses NLP to automate the extraction of 22 quality measures and integrates electronic reporting directly with CMS information systems**
- **Saves costs with remote RAC audits over networks, with automatic hyper-linking to the justifying document(s)**

Other Use Cases in Medicine for Zato Interoperability Platform

- Clinical Documentation Improvement
- Clinical/Genomic informatics: Indexing and analyzing in parallel patient genomic variant data with patient clinical data to correlate specific phenotypic expression with specific genomic variants across data centers
- Risk analytics for the insurance industry: integrating underwriting, claims, demographic data to mitigate risk
- Real time analysis for health crises such as pandemics, where data sets of structured and unstructured data can be spanned across networks
- Acceleration of research breakthroughs from greater data reach and productivity

Advantages of Semantics Assisted Medical Ontology

- Relatively accessible to both healthcare professionals and linguists, thus easing the task of development and maintenance

- Elegantly enables the addition of a semantic layer to syntactic parsers, thus easing the burden grammatical analysis

- Provides a measure of semantic distance between partial and fully formed concepts

- Maps well to both Information Extraction (IE) and Information Retrieval (IR) applications

Beyond HL-7: Indexed Natural Language Processing

- Leverages a comprehensive medical ontology and indexes massive data with an understanding of the relationships between anatomy, findings, disease, morphological abnormalities, procedures, symptoms, organ systems, medical tests and treatments, and ICD-9 And CPT-4 coding criteria, with path to transition to ICD-10
- Indexes data in its native location rather than consolidating data
- Normalizes a local view of data for presentation to the outside world while maintaining their local “dialect” internally
- Local data elements also map to data in other silos and into local “dialects”
- Automated granular extraction and information retrieval from structured and unstructured documents
- Codes or audits millions of records within hours, without load on host system

Zato Data Security, Tracking, and Accountability for Oversight, Compliance, Auditing, Comparison of Cost Effectiveness

- All Data and indexes remain where the data are created, stored, and protected
- All data access is subject to user role-based discretionary access control
- Optional download encryption and control under Digital Rights Management
- Security enforced and sharing permission controlled at each organization
- Server-to-server security; cross-network data encryption end-to-end
- All user activity is logged at the source for tracking, with global analysis of logs
- Successful application in the Intelligence Community processing Suspicious Activity Reports from US banks at large scale under the Bank Secrecy Act
- Successful cross-domain implementation controlling security across classified and unclassified data and networks in parallel from a unified interface

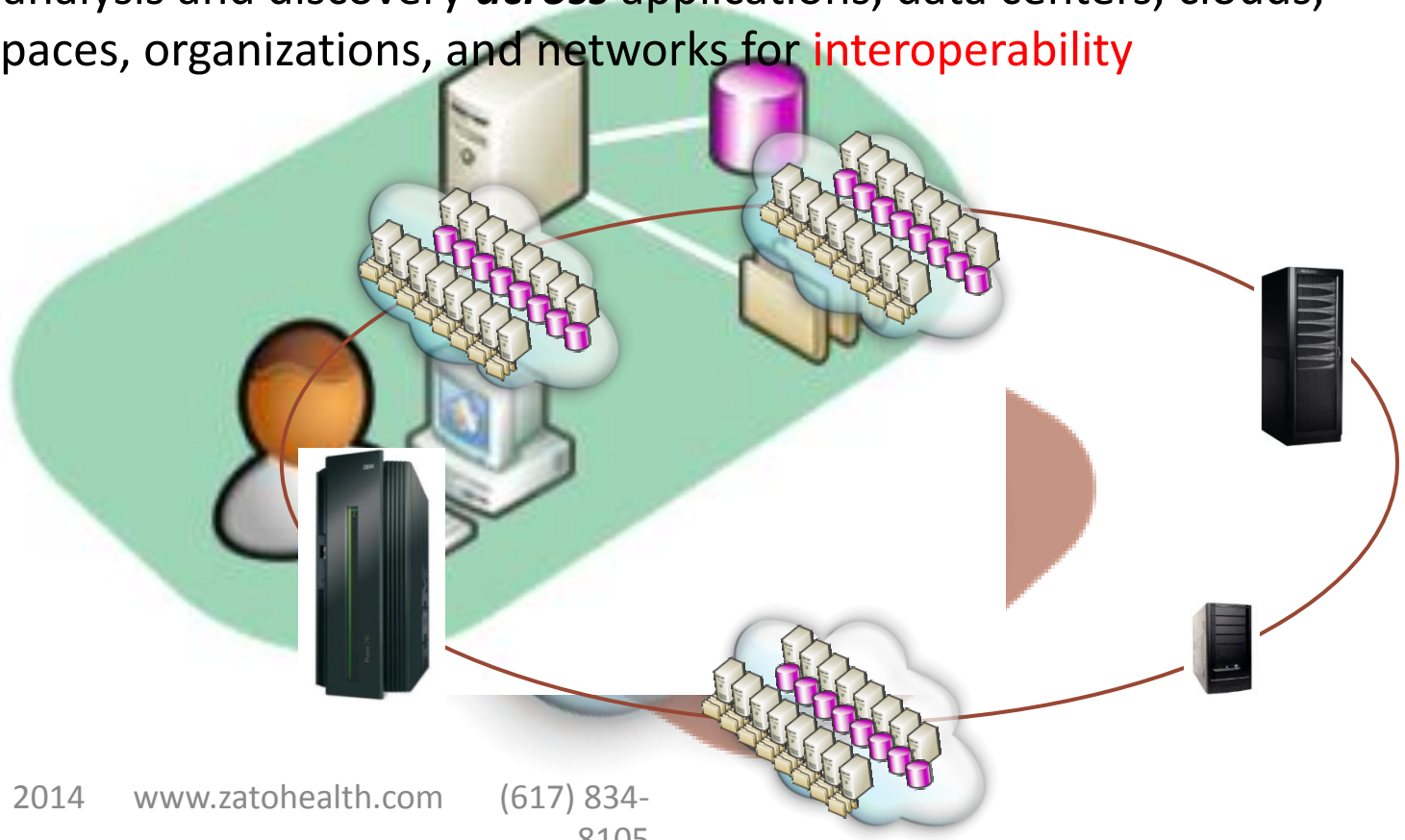
Zato Dynamic (On-the-Fly) Concept Recognition & Entity Extraction Across Decentralized Text-Rich Files, Tabular Data, and Streaming Data

- Addresses and Geo-Spatial Locations
- Security Classification Indicators
- Government Program Names
- Medical concept recognizers
- Person Name and Variants
- Drivers License Identifiers
- Topics and Noun Phrases
- Social Security Numbers
- Miscellaneous Numbers
- Domain Specific Lists
- New concepts in software update releases
- Credit Card Numbers
- Map Page Identifiers
- Special Name Lists
- Passport Numbers
- E-Mail Addresses
- Company Names
- Financing Terms
- Phone Numbers
- Glossary Terms
- Date of Birth
- Nicknames

A toolkit is also included for licensees to deploy additional domain specific concept recognizers and entity extractors

Evolution from Virtual Computing to Net-centric Information Fusion

- Virtual Machines (VMs) running in parallel on a single server increased efficiency
- Distributed processing server clusters within a 'Cloud' computing data center
- Secure analysis and discovery ***across*** applications, data centers, clouds, name spaces, organizations, and networks for interoperability



Maturity Levels – Where Innovation is Needed

• *Data Collection of* numeric, categorical, fielded ('structured') medical data and descriptive, context rich textual ('unstructured') patient data

• *Central Analysis of structured Data* in a EHR system or other database

• Exploitation of unstructured data within one EHR and data center

• Correlation across structured and unstructured data in one data center

• Interoperability: Unmet industry needs for effective, proactive data exploitation and correlation *across* combined structured and full text data in centralized and decentralized EHR data silos, other databases, and file systems across organizations, data center locations, and networks in parallel

Technical Barriers to Extending Reach of Analysis and Discovery

- Tabular analysis alone limits the kinds of questions that can be asked and the kinds of data navigation, feedback, and discovery that are possible

- Stovepipes destroy user productivity in the face of limited personnel budgets

- Driving need for effective *and affordable* technical solution for *decentralized* search, analysis, global views, global navigation, concept recognition, entity extraction, discovery 'on the fly' in parallel over multiple networks across text in file systems and structured tabular data and other application repositories held by different organizations at massive scales

Need to Supplement and Complement Tabular Database Processing

- All data for tabular analysis must be consolidated and aggregated centrally or replicated to each location – limiting views and discoveries *within* a data center
- Consolidated data centers and 'clouds' become larger disconnected silos
- Diversity and scale of data overwhelm a single centralized table space
- ETL of data to a central table space not required for > 90% of data
- Physical replication and distribution of patient specific data to multiple organizations increases security risk

Successful Prior Applications of the Zato Team with Information Fusion Component Technologies

(testimonials available from agency directors and program managers)

- National security intelligence at massive scale
- Cyber crimes analysis
- Discovery and interdiction of financial fraud (terrorist financing)
- Higher productivity in discovering threats to international pharmaceutical supply chain

"Techniques until recently available only to secret gov't agencies *will increasingly become mainstream in healthcare…."*

Underlying Components Field Proven in High Volume Uses in Applications for National Security

- > 10,000 intelligence community users with security clearances
- > 200,000 query and analysis actions executed daily
- > 15 billion database records and text-rich files indexed from over 100 different agency sources
- Highly secure cross-domain query and analysis used for discovery across multiple data centers of multiple agencies and multiple networks in parallel
- Fully distributed logging, tracking proven for compliance audits

Productivity Gains Reported from the Underlying Analysis and Discovery Software Component

"before 2002, it would take 32,222 hours to run 1,000 names and birth dates across 50 databases. Now agents can make such a search in 30 minutes or less"

(in articles by Washington Post, CNN, CBS Evening News from interviews with FBI officials - copies available)

Differentiation of Zato Component Technology

- Senate Armed Services Committee report of unmet needs for interoperability
- FY 12 DOD Budget Authorization directed a DOD study and cross-agency pilot
- Requirement included capability for interoperability to scale across the data repositories of the entire Defense Intelligence Information Enterprise (DI2E)
- No existing capability to effectively satisfy all three pivotal requirements:
  - **Cross-domain secure access simultaneously to data on multiple networks**
  - **Enables discovery of relationships among data in different data centers**
  - **Standards-based API enabling ease of integration for interoperability across networks and data centers**

Government Contractor's Scalability Estimate for Decentralized Analysis and Discovery at Scale

- Trillions of records and documents
- Hundreds of thousands of authorized users
- 6 Intelligence processes
- 7 intelligence disciplines
- 6 Intelligence product categories
- 12 agencies with many data sources in each
- 3 data OPS levels
- 5 primary user locations
- 4 security levels

Zato's Underlying Information Fusion Component Addresses DOD Requirements* for Cross-Agency Interoperability

(a) Plan (Independent Government Team)

- Assessment of all the current and planned advanced query and correlation systems
- Determine where duplication can be eliminated,

(b) Pilot (Contractor Team)

- Demonstrate an enterprise query and correlation capability through the DI2E program to achieve
  - Conduct **complex, simultaneous queries** by a large number of users and analysts **across numerous, large distributed data stores** with response times measured **in seconds**.
  - **Scale up to operate effectively on all data holdings of the Defense Intelligence Information Environment (DI2E)**
  - Operate across **multiple levels of security** with data guards.
  - Operate effectively on both **unstructured data and structured data**.
  - **Protect sources and methods**, privacy, or both.
  - **Control access to data** by means of on-line electronic user credentials, profiles, and authentication
  - **Initial Period of Performance** through September 30, 2014.

(c) Report (Government Team with inputs from Contractor Team)

- November, 2012. The report shall set forth the plan developed under subsection (a) and a description and assessment of the pilot program conducted under subsection (b).

*** Assessment of dozens of commercial, government developed, and open source software alternatives**

Contact Information for Questions

Paul McOwen, President and CEO

pmcowen@zatohealth.com

617 834-8105